Enclosure C

The Centers for Medicare & Medicaid Services (CMS) conducted an extensive consultation process in reaching the decisions on risk adjustment. We held telephone consultations with over 20 M+C organizations and several trade associations to determine possible approaches to burden reduction and to solicit each organization's opinions on possible changes to the risk adjustment model. CMS convened a public meeting where we provided M+C organizations and other attendees with a description of three risk adjustment model types: limited diseases model (incorporating only 6 disease groups), a selected significant diseases model (with an example of 25 conditions provided), and an all-significant disease model (incorporating 86 conditions[1]). CMS requested comments from the public on risk adjustment model types, including site neutrality and the number of conditions that should be included in the future risk adjustment model, as well as burden reduction efforts.

CMS received 24 sets of comments on the risk adjustment model. The sources of these comments were: 9 sets of comments came from M+C organizations; 7 sets came from managed care and provider trade associations and related organizations; 4 sets came from two risk adjustment model developers; and 4 sets came from health-related organizations. CMS indicated at the public meeting that a decision would be made on the specific risk adjustment model based on a number of criteria, including conceptual understanding to clinicians, providers, and plans, comparative analytic performance, incentives, and appropriateness for payment. The comments and CMS's response are discussed below as well as CMS's decisions on site neutrality, the number of conditions and the type of model. Comments on other issues will be considered as those decisions are made over the next year.

Comment: Site Neutrality

Of the 24 commenters, 14 universally supported site neutrality for the risk adjustment model. The remaining 10 from whom comments were received did not comment specifically on this issue. Many of the commenters favoring site neutrality said that a site neutral model recognizes managed care's increasing use of outpatient settings for treatment and management of their population rather than an inpatient setting. It allows "providers the flexibility to provide whatever combination of care is most effective, in the most appropriate delivery setting, and in the least restrictive environment possible."

One commenter supported an "all inpatient plus model." An "all inpatient plus" model is one in which all diagnoses from the inpatient setting, including secondary diagnoses, are included together with a limited set of ambulatory diagnoses. Commenters expressed concern that a model that differentiates payment according to inpatient and outpatient settings, such as the all inpatient plus model, may create incentives to hospitalize. "If a model were selected that included all inpatient diagnoses but only select outpatient diagnoses, the incentive to hospitalize marginal cases would be enormous." "The 'All Inpatient Plus' approaches fail to eliminate the incentives to: 1) admit patients for short hospital stays to take advantage of increased payments; and 2) recruit or avoid membership based on health conditions included or excluded."

[1] Note that the January 16, 2002 public meeting materials erroneously labeled this model as 100 conditions.

CMS Response

CMS agrees that site neutrality is the preferred approach to making risk-adjusted payments to M+C organizations as it recognizes managed care's increased use of outpatient settings for the provision of care. CMS also believes that models that make differential payments for a hospitalization may provide incentives to hospitalize. A site neutral model avoids those incentives as payments are not differentiated by the site of care in which the diagnosis was received.

Comment: Number of Conditions in the Model

CMS specifically asked for comments on three types of models: a selected significant disease model with only 6 conditions; a selected significant disease model (with an example of 25 conditions); and all significant disease model with 86 conditions.

Ten commenters provided no specific comments on the number of conditions to include in the model. One stated they would need more information before being able to make such a determination.

One commenter wanted a model that used 6 conditions in addition to all inpatient diagnoses. They preferred a short list due to their estimates of the burden of data collection. "The difference in [the data collection] effort for six condition groups (estimated only to have 25 ICD-9s) and 100 (estimated to have 1000 ICD-9s) will be gigantic." They cited increased burden associated with collecting data for a model with more conditions including the need to hire 50 additional employees at a cost of $2.5 million a year, if a model larger than 6 conditions were chosen.

One commenter preferred adding no more than 25 conditions to an inpatient model. They suggested adding outpatient diagnoses that add predictive power to the model while reducing the data burden. They favored building upon the inpatient model because "M+C organizations already have spent a significant amount of resources in the development of processes and systems to support the submission of inpatient encounter data, and a significant amount of medical costs are associated with this site of service."

One favored the number of disease conditions to be between 6 and 25 disease groups where the diagnoses could come from either an inpatient or outpatient setting. The reason for this number of disease groups was related to the potential financial impacts if M+C organizations were not able to collect sufficient data. "If M+C organizations are unable to collect a sufficient amount of the necessary data, the negative impact on revenues will be significant, and a fundamental purpose of health status-based risk adjustment, establishing a closer relationship between payment and the cost of providing care to M+C enrollees, will be seriously undermined."

Two commenters said that they would favor a "comprehensive" model in the long term if the phase-in of risk adjustment beginning in 2004 were slowed. One reason for a slowed implementation of a "comprehensive" model according to one commenter is because they and "most other M+C plans cannot meet the data requirements of such a system at this time, and payments resulting from such a system would be inappropriate."

Nine commenters said they wanted the most comprehensive of the models or the "all-significant disease" model. One reason for supporting an all-significant disease model includes better predictive power of the model. "Maximizing the predictability of the risk adjuster is essential to providing a credible adjustment to payment. The use of the full 100 diagnoses applied neutrally fulfills this goal." Another reason given for selecting an all-significant disease model is that it reflects more of the chronic conditions that affect a Medicare population.

Several commenters expressed an interest in going beyond a comprehensive model to recognize the special needs and costs associated with the care of a frail population. These limited models do not adjust payments for some of the most common conditions that the frail elderly suffer, such as pneumonia, dehydration, depression, and other conditions." "Limiting the risk adjustment models to an arbitrary number of diseases does not adequately adjust payments for the highest-expenditure, chronically ill individuals." Further, an all-significant disease model is preferred because it emphasizes managing a full range of diseases, not just a limited set. "A model with fewer condition categories would penalize our work to avoid such deterioration in health for a beneficiary, who still may have a serious condition that is not considered by a model with fewer categories."

Finally, several commenters suggested that CMS use certain criteria to select the diagnoses for the model. One provided a list of diagnoses and concerns about those diagnoses. For example, they suggested that certain frequently miscoded diagnoses be eliminated from the list and that some diagnoses be added because they are increasingly associated with high cost care.

CMS Response

Three sets of comments support a model with 25 conditions or less. There are a number of problems with models that include 25 or fewer conditions. Models in this range do not improve the payment accuracy over the current PIP-DCG model, which is the goal of adding ambulatory diagnoses. In addition, a 6-condition model would continue overpaying M+C organizations for their lowest cost enrollees by about 85 percent.

CMS believes that models in the range of 6 to 25 conditions are not defensible clinically. For example, a 25-condition model would exclude many clinically important condition categories groupings such as breast cancer, hip fracture, and alcohol and drug dependence.
Payments for beneficiaries with those diseases would likely be too low.

Commenters supporting fewer condition models also cite the burden associated with submitting the data needed for a larger model. However, one of the first steps CMS took in redesigning the risk adjustment model for 2004 was to reduce data burden to M+C organizations. Some of these reductions include less frequent submission, the use of simplified formats, and collecting only 5 data elements. All told, these system efficiencies result in more than a 95 percent reduction in data submission burden over the previous encounter data approach. CMS achieved additional burden reduction by reducing the number of conditions in the model from 86 (in the all-significant disease model) to about 61 in the selected significant disease model. Using less diseases results in a 98 percent data submission burden reduction. All of the data burden reduction efforts have been widely supported and endorsed by the M+C industry. Commenters were particularly supportive of simplified reporting of fewer data elements and edits, the flexibility in reporting format options and the reduced volume of transactions.

The number of conditions contained in the approximately 61 disease group model represents a small or incremental increase in data submission effort from the fewer condition models. This is because M+C organizations need only to report the selected diagnoses for which a beneficiary is diagnosed and that CMS achieved additional burden reduction by reducing the number of conditions in the model from 86 (in the all-significant disease model) to about 61 in the selected significant disease model are relevant for payment. Further, M+C organizations need only report the diagnoses once during the data reporting period. Therefore, we believe that citing data burdens as an obstacle to using a larger model is not in alignment with actual data submission burden reduction already achieved by CMS.

In response to the comment that models with more than 25 disease groups are so burdensome that M+C organizations will be unable to submit adequate data to ensure that their revenues are not adversely impacted, CMS believes that this argument is not supported by the information we have previously received. Most claims-based M+C organizations suggested in phone consultations that they prefer to submit all data because their physicians and other providers already comply with submitting their claims data in order to receive payment. Therefore, claims-based M+C organizations should not face difficulties in submitting their data. M+C organizations that do not receive claims or encounters from their physicians may use a superbill to collect diagnoses for risk adjustment. Superbills permit the use of collapsed diagnosis codes.

Regarding the comments related to the burden of hiring additional staff and the associated high costs if a larger model were chosen, CMS is aware that there are staffing and cost concerns associated with risk adjustment data collection from physicians and providers. We believe that we have addressed those concerns by reducing the number of diagnoses to be submitted and the frequency with which data is submitted. Also, the use of a superbill should assist M+C organizations with their data collection efforts.

The majority of organizations that commented on the size of the model support an

"all-significant disease model," with one supporting it only after a transitional period using a smaller model. Another suggested a slower phase-in of risk adjustment. The risk adjustment phase-in is mandated by the Benefits Improvement Protection Act of 2000 and a change in the phase-in schedule would require legislation.
CMS agrees that an "all-significant disease" model improves the predictive power of the model and thus improves payment accuracy. However, CMS's research on the models also shows that a model with all-significant diseases improves the predictive power of the model only slightly over the "selected significant disease" models. In an effort to require the minimum data that need to be submitted while balancing improvements in payment accuracy and providing incentives for plans to provide disease management services for certain serious chronic conditions, CMS has selected a model with approximately 61 disease groups.

This selected significant disease model was determined by reviewing the diagnoses in the all-significant disease model and excluding those disease groups where the Medicare prevalence was low or the predictive cost was low. The selected significant disease model also substantially reduced the number of ICD-9 codes that need to be collected by more than 60 percent beyond the all-significant disease model. In addition, clinical criteria were applied. Clinicians we consulted concluded that models smaller than 61 conditions would exclude too many diseases. For example, models in the range of approximately 61 disease groups include important conditions such as breast cancer, hip fracture, and alcohol and drug dependence that are not included in the smaller models. Other criteria considered by the clinicians and CMS staff included incorporating frailty diagnoses, such as protein calorie malnutrition, decubitus ulcer of the skin, and major depressive disorders, and eliminating diagnoses that might result in gaming or which are frequently miscoded. Finally, CMS recognizes that M+C organizations, particularly claims-based organizations, may prefer to submit all diagnoses. The selection of a selected significant disease model will not preclude M+C organizations from continuing to submit all diagnoses.

Comment: Type of Risk Adjustment Model

Of the 24 commenters, 5 addressed desirable features that the type of risk adjustment model selected should contain. Several of the model developers suggested that selected significant disease versions of their models performed almost as well as all-significant disease versions. One developer suggested the importance of a stable, clinically determined model that could be used for multiple purposes including clinical management. Another organization stressed the importance of updating and recalibrating the model as coding practices and treatment costs, and patterns change.

CMS Response

CMS worked with 5 developers to produce models over the last several years: one was developed by Health Economics Research (HCC), one by Richard Kronick of University of California, San Diego (CDPS), two by a team from Johns Hopkins University (ADG, ACG), one by a team from RAND (CDRISC), and one by 3M Health Information Systems (CRG). Some of

the models are "additive" meaning that individual conditions contribute to an increased payment as they are added. Other models are "categorical," meaning that the payment is based on the overall health status category in which an individual is placed. Two of the models are categorical and the rest are additive.

CMS agrees with the comment that a selected significant disease model performs almost as well statistically as an all-significant disease model at predicting costs for the following year. The predictive performance, as measured by the r^2 statistic, did not vary greatly across the models under consideration. Another measure of performance, the predictive ratio statistic, is used to test how well the models predict for atypical groups of people ("biased groups"). Predictive ratios look at the predicted costs for atypical groups divided by the actual costs for those same persons. Predictive ratios in the range of 1.0 are accurate while those below 1 pay too little and above 1 pay too much. Again, most of the models in the range of 25 to 86 disease groups would perform acceptably for most biased groups. CMS believes that most of the developers with whom we worked could modify their models to incorporate a smaller number of diagnoses than were used with their original models.

While many of the models would perform adequately from a statistical perspective, other criteria were considered in selecting the specific model, including sensitivity to coding practices, ease of understandability/transparency, and ease of modification/timeliness. Among the additive model structures, the ADG model clusters diseases by characteristics such as time-limited, likely to recur, likely to result in hospital stay, chronic, etc. Costs are associated with these groupings. The specific diseases are not explicitly visible. The CD-RISC model has disease categories that are hierarchical within body systems. The ranking of a disease in a hierarchy is determined by the code's inherent severity (and costliness) within the body system and by the presence of comorbidities in other body systems. In contrast, the HCC and CDPS additive models are very simple, containing disease groups and combinations of diseases. There are cost hierarchies within disease types, such as cancers, gastrointestinal conditions, and cardiac conditions. In CMS' discussions with plans, policymakers, and clinicians, we found that the additive models were very easy to explain and understand. This was particularly apparent during our discussions about the diseases to be included or excluded in a selected significant disease model. The frequency of each disease group in the population and the associated costs were readily apparent. We could easily re-estimate the model and determine the effect of removing a disease group without concern over the unexpected consequences. This ease of modification is not present with the categorical models. Of the additive models, the CMS clinicians preferred the assignment of ICD-9 codes to the HCC disease groups rather than the CDPS disease groups.

Concerning the importance of a stable clinically determined model that could be used for multiple purposes including clinical management, we have described our process of clinical review of the model in the earlier comment section. The HCC model was reviewed by physicians during its development as well as by CMS staff as we deleted conditions from the all-significant disease model. CMS has chosen the simplest model clinically, given that payment is the primary purpose, but at the same time believes that this model can be used by physicians for clinical management purposes.

On the final comment related to the selection of the type of model, CMS agrees with the importance of updating and recalibrating the model as coding practices and treatment costs and patterns change. As mentioned above, the ease of modification of the model was a key criterion in selecting the type of model. The additive models lent themselves more readily to modification. CMS can easily re-estimate the model and determine the effect of removing a disease group without concern over the unexpected consequences.

While not directly addressed by the commenters, CMS also considered several other criteria in selecting the specific model:

- In selecting an additive rather than a categorical model, we considered whether an additive model is more susceptible to more accurate coding than a categorical model. One reason to believe that the additive models may be more susceptible to upcoding or more accurate coding is that the rewards are readily transparent. For each additional coded diagnosis for one of the selected diseases, a plan would receive an extra payment. However, the categorical models also reward the same coding practices although not as directly. In the categorical models, additional diseases may advance an individual into a higher severity category or to a category which include two chronic conditions rather than one. On the other hand, the additive models are more easily monitored for auditing purposes. Monitoring and auditing attention may be directed to plans that have unusually high levels of specific diseases.

- Ease with which M+C organizations and others can understand the model and determine an enrollee's risk score. CMS's experience with the current inpatient model has shown that M+C organizations have varying levels of ability to calculate risk scores and understand the model. The additive models are more easily understood and will be easier for M+C organizations and the public to understand in calculating risk scores.

- The accessibility of the software for use by the public, including M+C organizations. Software that is developed entirely with U.S. government support will not require negotiations over copyrights. It is essential that all of the details and an accompanying public version of a grouper for the selected model be made available to M+C organizations and the public.

Based on all of the criteria above, we believe that the CMS-modified version of the HCC model (CMS/HCC) is the best model to be used for M+C risk adjustment. It is highly accurate, easy to understand, easy to modify and recalibrate, and solely developed with government funds. It also can be implemented in the timeframes we are facing to implement risk adjustment according to statutory requirements.